ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

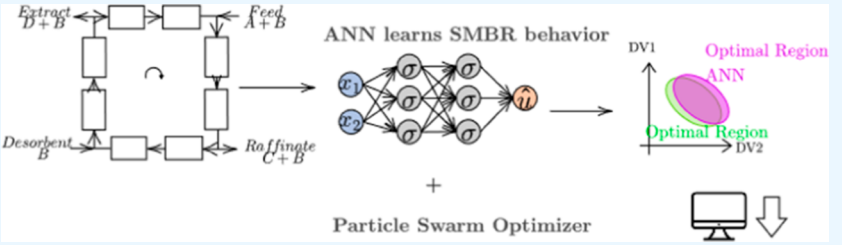
# Novel Framework for Simulated Moving Bed Reactor Optimization Based on Deep Neural Network Models and Metaheuristic Optimizers: An Approach with Optimality Guarantee

Vinícius V. Santana, Márcio A. F. Martins,* José M. Loureiro, Ana M. Ribeiro, Luana P. Queiroz, Carine M. Rebello, Alírio E. Rodrigues, and Idelfonso B. R. Nogueira*

Cite This: *ACS Omega* 2023, 8, 6463−6475

Read Online

ACCESS | Metrics & More | Article Recommendations




**ABSTRACT:** Model-based optimization of simulated moving bed reactors (SMBRs) requires efficient solvers and significant computational power. Over the past years, surrogate models have been considered for such computationally demanding optimization problems. In this sense, artificial neural networks—ANNs—have found applications for modeling the simulated moving bed (SMB) unit but not yet been reported for the reactive SMB (SMBR). Despite ANNs' high accuracy, it is essential to assess its capacity to represent the optimization landscape well. However, a consistent method for optimality assessment using surrogate models is still an open issue in the literature. As such, two main contributions can be highlighted: the SMBR optimization based on deep recurrent neural networks (DRNNs) and the characterization of the feasible operation region. This is done by recycling the data points from a metaheuristic technique—optimality assessment. The results demonstrate that the DRNN-based optimization can address such complex optimization while meeting optimality.

## 1. INTRODUCTION

The simulated moving bed reactor (SMBR) extends the simulated moving bed (SMB) process, where the continuous chromatographic separation enhances chemical reactions. The SMBR has attracted significant attention in the past years,[1,2] especially for synthesizing oxygenated compounds and isomerization reactions.[3,4]

Model-based optimization of SMBR units implies the simultaneous solution of dynamic nonlinear partial differential equations, requiring efficient solvers and significant computational power. Moreover, several other issues arise in optimizing SMBRs due to the coupling of reaction and separation. It reduces the degree of freedom;[5] that is, fewer variables can be manipulated to achieve purity and conversion requirements and conflicting objectives; namely, improving one performance indicator worsens the others in a nontrivial way.

Therefore, SMBR optimization is not a trivial task that motivated many studies to address it with mixed approaches—first-principles models with varying simplifications and deterministic/heuristic optimization methods. Tie et al.[6] proposed an epsilon-constrained multiobjective optimization with a full-discretization model to maximize the production rate of propylene glycol methyl ether acetate and conversion of ethyl acetate solved with interior point OPTimization techniques. Ray and Ray[7] proposed a multiobjective optimization using nonsorted genetic algorithm II (NSGA II) to maximize yield and purity in biodiesel production. Subramani et al.[8] presented a multiobjective optimization problem for the methyl tertiary butyl ether synthesis in a Varicol SMBR optimized with the NSGA algorithm. Nogueira et al.[9] proposed a single objective optimization problem for producing *n*-propyl-propionate with true moving bed reactor approximation using a particle swarm optimization (PSO) method.

The optimizer solves the SMBR model several times to evaluate the objective function during optimization. The simulation time can grow significantly depending on model

Received: October 19, 2022
Accepted: January 19, 2023
Published: February 7, 2023

simplifications—reaction rate law, equilibrium equations, number of columns per zone, and nonisothermal operation. For the design of the SMBR, the optimization time may not be prohibitively high. However, the high-fidelity model may become an unfeasible option in an actual plant coupled with up- and downstream equipment or in real-time applications. Simplifications of the high-fidelity model are often employed for SMB separation by using the true moving bed (TMB) concept and the SMBR with the true moving bed reactor (TMBR). However, it has been shown that this approximation can fail to describe the SMB and SMBR in several scenarios.[10,11]

Over the past years, reduced-order/surrogate models have been considered for optimization problems. Surrogate models are mathematical models identified using statistical techniques and attenuating the associated computational costs of an optimization problem. In this sense, artificial neural networks—ANNs—have found applications for these systems for modeling the SMB unit (without reaction).[12−14] However, they have not yet been reported for the reactive SMB (SMBR). ANNs are powerful models composed of multiple processing layers that learn representations of data with various levels of abstraction[15] and are proven to be able to approximate any nonlinear $C_1$ continuous functions[16] and overcome the curse of dimensionality.

Despite ANNs' high accuracy for nonlinear function approximation, assessing their converges and capacity to represent the optimization landscape competently is essential.[17] To this end, it is important to use strategies to identify high-accuracy surrogates (low bias) and investigate the optimality, that is, convergence to the true global optimum. However, a consistent method for optimality assessment using surrogate models is still an open issue in the literature. Optimality is usually assessed by visual inspection and pointwise (i.e., by comparing the high-fidelity model-based optimal point with the surrogate model-based optimal one).

A few works have recently shown one approach to evaluating the optimization results with a cluster of points instead of a single evaluation. Park (2013) proposed a bootstrap approach, and Nogueira et al.[18] and Rebello et al.[19] used recycled data from PSO—however, using the high-fidelity model. As shown in refs 9 and 19, it is possible to map the optimal region through the data population generated from the PSO algorithm. As such, the method allows using an available massive database from PSO to produce meaningful information about the optimal solution and increase its robustness level by evaluating its associated uncertainty. Addressing this issue is one main contribution of this work and still an open issue in the literature.

Hence, two main contributions of this work can be highlighted: the SMBR optimization based on deep neural networks (DNNs) and the challenging characterization of the feasible operation region (FOR). Therefore, the DRNN-based optimization is assessed in the optimization of an SMBR unit. This is done by recycling PSO data points, allowing the optimality assessment. A case study presents the synthesis of *n*-propyl-propionate in an SMBR-4 unit. The present work proposes a novel framework for SMBR units' PSO-oriented optimization based on DNN models.

## 2. METHODOLOGY

**2.1. SMBR Model.** The SMBR comprises a set of fixed-bed columns arranged in a recirculation loop. There are two inlets (feed and desorbent) and outlet streams (extract and raffinate). The packing material(s) have either adsorptive or catalytic properties, and the feed stream contains the reactant(s). The products are separated from the reactants as the reaction occurs within the unit. Figure 1 illustrates a four-

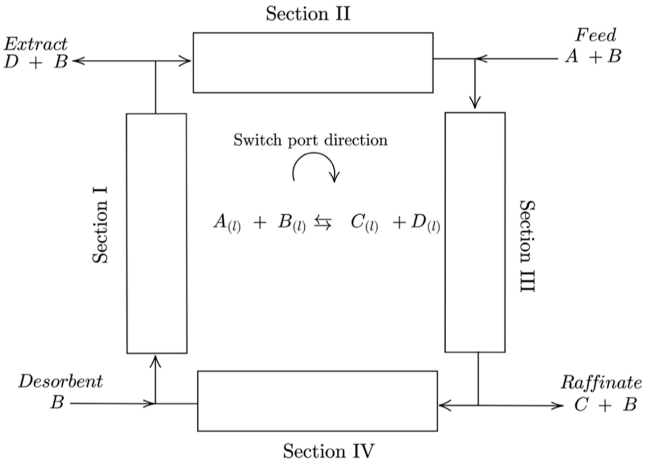



**Figure 1.** Diagram of a SMBR unit.

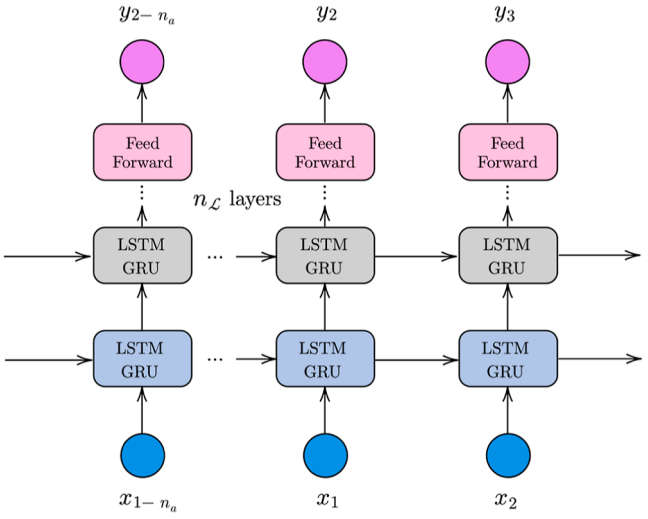



**Figure 2.** Proposed stacked RNN model.

zone SMBR for a reversible bimolecular hypothetical reaction. For such cases, it is usual that one of the reactants is used as a desorbent. The reactants (A + B) are fed between sections II and III; the raffinate stream is located between sections III and IV. It is where the less adsorbed product (C) is collected, the desorbent stream is located between sections I and IV, and the extract stream lies between sections I and II from which the more adsorbed product (D) is withdrawn.

In the present work, the SMBR-4 model for the synthesis of *n*-propyl propionate proposed in ref 11 was implemented in gPROMs[20] to serve as the high-fidelity model to be optimized and source of data for the DRNN. In this case, propanoic acid, propanol, *n*-propyl propionate, and water play the A, B, C, and D roles, respectively, considering Figure 1 as a reference. The model equations and parameters are displayed in Tables A1 and A2. The parameters were previously estimated by Nogueira et al.[9]



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

Table 1. Input and Output Signals' Ranges

| variable | minimum | maximum |
|---|---|---|
| Input Signals (Designed) | | |
| $C_F^{ProAc}$/(mol/L) | $1.40 \times 10^{-1}$ | $1.33 \times 10^{1}$ |
| $Q_D$/(dm$^3$/min) | $1.00 \times 10^{-3}$ | $4.99 \times 10^{-3}$ |
| $t_s$/(min) | $1.00 \times 10^{1}$ | $5.99 \times 10^{1}$ |
| Output Signals | | |
| raffinate purity/(%) | 0.65 | 99.99 |
| conversion/(%) | 1.28 | 99.97 |
| productivity/(mol/L/day) | $3.432 \times 10^{-5}$ | 16.74 |
| desorbent consumption (L/mol) | $4.785 \times 10^{-1}$ | 89.95 |

Table 2. Hyperparameter Search Space for the RNN

| hyperparameters | search space |
|---|---|
| initial learning rate | {$1 \times 10^{-4}$, $3.162 \times 10^{-4}$, $1 \times 10^{-3}$} |
| batch size | {4, 8, 16} |
| number of recurrent layers | {1, 2, 3} |
| recurrent layer type | {GRU, LSTM} |
| number of neurons in the recurrent layers | {100, 180} |
| activation function in the recurrent layers | {relu, tanh} |
| number of neurons in the intermediate fully connected layer | {20, 60} |
| activation function in the fully connected layer | {relu, tanh} |

A total of four performance indicators of separation and reaction were considered: raffinate purity, limiting reactant conversion, productivity, and desorbent consumption. Table A1 shows the equations for raffinate purity, conversion, desorbent consumption, and productivity in the case study of propyl-propionate synthesis.

**2.2. DNN Model—Building and Estimation.** Tables A1 and A2 show that the SMBR is an inherently dynamic process. Hence, using the discrete-time recurrent neural network (RNN) model is the most appropriate way of introducing inductive bias and shortening the amount of training data needed for long-horizon predictions (simulation mode). The difference between prediction and simulation is well discussed elsewhere.[14] Thus, this architecture is proposed as a surrogate model of SMBR-4. Several RNNs, one for each performance indicator (time-averaged raffinate purity, conversion, desorbent consumption, and productivity), were identified, that is, multiple input single output (MISO). Here, we propose a flexible approach with a mixed stack of state-of-the-art RNN architectures built—long−short-term memory (LSTM) or gated recurrent unit (GRU) and feed forward layers. Figure 2 illustrates the proposed approach.

Table 4. Rescaled Decision Vectors at the Optimal Point for Rigorous and DRNN Models (Rescaled to the Original Units of Measurements)

| model | $Q_D/\left(\frac{dm^3}{min}\right)$ | $C_F^{ProAc}$/(mol/L) | $t_s$/(min) | objective function |
|---|---|---|---|---|
| DRNN | $1.68 \times 10^{-3}$ | 13.33 | 47.76 | −1.17 |
| rigorous | $1.77 \times 10^{-3}$ | 13.33 | 47.51 | −1.14 |

Once the general model architecture is proposed, the identification procedure, that is, a parameter estimation strategy, has to be developed, which includes the design of experiments (DoE) or data acquisition, time series preprocessing, and the parameter estimation itself, which involves defining the neural network topology, cost function and training policy—training set size, optimizer type, and parameters.

Among all possible input variables in the SMBR, the ones with the highest impact on its performance indicators were chosen, according to a strategy developed by Santana et al.[11]—the concentration of the limiting reactant in the feed stream (propanoic acid), the desorbent flow rate, and the switching time. Then, the DoE was performed using the Latin Hypercube Sampling (LHS) algorithm[21] to sample the input space ($\mathbb{R}^3$) with minimal cross-correlation between variables. The sampled inputs were fed to the high-fidelity model simulator as a sequence of step signals persisting until the system reached the cyclic steady state. The simulator outputs (performance indicators) were collected and stored. The LHS is one among many options available for DoE, which includes the popular Sobol sequences.[22] Each method has its own advantages and disadvantages. The literature in ANN-based surrogates presents relevant usages of both methods.[19,23−28]

The available time series preparation into "experiments", often named "time windows", forces the model to learn the whole response to a step change in the input variables. In simple terms, a set of small chunks of time series are created, wherein each experiment (chunk) $E_i$ consists of a set of tuples $\{(\boldsymbol{u_1} \rightarrow y_2), (\boldsymbol{u_2} \rightarrow y_3), ..., (\boldsymbol{u_3} \rightarrow y_{n_i})\}$.

The time series preprocessing involves identifying long-term dependencies (order) in the time series data. In the present work, the dynamic system is modeled as a nonlinear state space representation as described in ref 29 It implies that for a given

Table 3. Results of Best Hyperparameters for Each Performance Indicator for the RNN

| hyperparameters | productivity | desorbent consumption | purity | conversion |
|---|---|---|---|---|
| initial learning rate | $1 \times 10^{-3}$ | $3.162 \times 10^{-4}$ | $1 \times 10^{-3}$ | $1 \times 10^{-3}$ |
| batch size | 4 | 4 | 4 | 4 |
| number of recurrent layers | 3 | 2 | 3 | 3 |
| recurrent layer type | {layer 1: LSTM, layer 2: LSTM, layer 3: LSTM} | {layer 1: GRU, layer 2: GRU} | {layer 1: GRU, layer 2: GRU, layer 3: GRU} | {layer 1: LSTM, layer 2: GRU, layer 3: GRU} |
| number of neurons in the recurrent layers | {layer 1: 180, layer 2: 180, layer 3: 180} | {layer 1: 180, layer 2: 100} | {layer 1: 100, layer 2: 100, layer 3: 180} | {layer 1: 180, layer 2: 100, layer 3: 180} |
| activation function in the recurrent layers | {layer 1: relu, layer 2: relu, layer 3: relu} | {layer 1: relu, layer 2: relu} | {layer 1: relu, layer 2: tanh, layer 3: relu} | {layer 1: relu, layer 2: relu, layer 3: tanh} |
| number of neurons in the intermediate fully connected layer | 20 | 60 | 60 | 20 |
| activation function in the intermediate fully connected layer | Relu | relu | relu | Tanh |



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

Table 5. Performance Indicators at the Optimal Point for the Rigorous Model and RNNs

| model | purity/(%) | conversion/(%) | desorbent consumption/(L/mol) | productivity/(mol/L/day) |
|---|---|---|---|---|
| DRNN | 99.58 | 93.10 | 1.18 | 3.56 |
| rigorous | 97.32 | 90.86 | 1.11 | 3.69 |

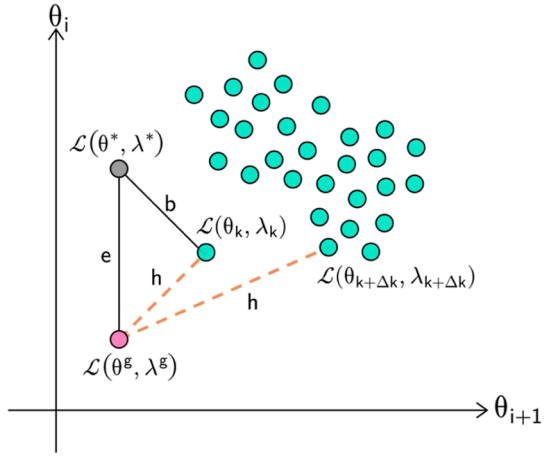



Figure 3. 2D geometrical illustration for confidence region derivation.

discrete-time dynamic system with white noise $v(k)$, noisy observations y depend on exogenous inputs $u$ and past estimated states $z$. It can be written

$$\boldsymbol{z}(k) = \boldsymbol{F}([\boldsymbol{z}(k-1), ..., \boldsymbol{z}(k-1-n_a), \boldsymbol{u}(k-d), ..., \boldsymbol{u}(k-d-n_b)])$$

$$\boldsymbol{y}(k) = \boldsymbol{z}(k) + v(k)$$

where $n_a$ and $n_b$ are the number of past values and $d$ is the delay. The system order ($n_a$ and $n_b$) is independent of the chosen function used to approximate the true unknown $F$ and should be carried before any parameter estimation.[14,30] To identify $n_a$ and $n_b$, the Lipschitz coefficient[31] method is used.

The neural network topology, cost function, and training policy are known as hyperparameters. They define the optimization problem and must be determined beforehand, that is, before the RNNs' weights and biases are estimated. The hyperparameter space comprises a set of both discrete and continuous variables, which makes their selection a complex task. In this sense, the state-of-the-art HYPERBAND[32] optimization algorithm is used here as it wisely allocates resources for training the most promising configurations. The search space for the hyperband comprised continuous and discrete variables: type of each recurrent layer, number of stacked layers, number of neurons per layer, activation function type, learning rate, and minibatch size. Despite using hyperband, other efficient automatic hyperparameter selection methods exist, such as multiobjective optimization problem training loop[33−35] and TRANSFORM.[36] At each run of HYPERBAND, the RNNs were trained with least-squares loss function and adaptive moment estimation (Adam) with TensorFlow default parameterization. After selecting the best configuration, the final architecture was trained longer with early stopping. Neural network building, training process, and hyperparameter tuning were implemented in TensorFlow 2.5.

**2.3. Optimization Problem Formulation.** The optimization problem is framed as a single objective function by weighting the SMBR performance indicators constrained by the model equations and decision variables' limits.

The objective function, evaluated at a cyclic steady state, and constraints can be written as

$$\min_x F_{obj} = \text{Prod} \times w_{\text{Prod}} + \text{Conv} \times w_{\text{Conv}} + \text{Dc} \times w_{\text{Dc}} + w_{\text{Pur}} \times \delta_{\text{Pur}}$$

$$[\text{Prod, conv, Dc, Pur}] = f(\boldsymbol{x})$$

$$\delta_{\text{Pur}} = 1 \text{ if } \text{Pur}(k) < \text{Pur}_{\text{min}}$$

$$\delta_{\text{Pur}} = 0 \text{ if } \text{Pur}(k) \geq \text{Pur}_{\text{min}}$$

$$\boldsymbol{x}_{\text{min}} \leq \boldsymbol{x} \leq \boldsymbol{x}_{\text{max}} \quad (1)$$

where Prod, Conv, Dc, and Pur are time-averaged productivity, conversion, desorbent consumption, and purity, respectively, and $x$ is the decision vector that contains the input variables values. $w_i$ are real values that weight each performance indicator; $\delta_{pur}$ is an indicator function that penalizes the objective function with $w_{pur}$ if the purity is lower than $\text{Pur}_{\text{min}}$. $f(x)$ is the model (mechanistic or DRNN) that maps the decision vector to the performance indicators. It can be either the rigorous SMBR-4 model or the RNN.

As shown in the previous section, the SMBR has four main performance indicators usually taken into account: productivity, conversion, purity of the product of interest, and desorbent consumption. The first three are usually desired to be as high as possible and the last as low as possible. It means that the corresponding weights in the objective function should reflect it; that is, the performance indicators to be maximized need a negative weight and vice versa. An in-house implementation of the global PSO method[37] is used to solve this optimization problem.

PSO was used as the optimization method as it has been demonstrated as an efficient and suitable approach for characterizing confidence regions in optimization problems.[19,38−40]

**2.4. Feasible Operating Region around the Optimal Point.** The strategy for determining the feasible operating region around the optimal point here used was proposed by Nogueira et al.[18] The uncertainty map is the denominated feasible operating region (FOR) in the referred work. The FOR is a subdomain of process operating variables where the process can attain a defined performance with a certain confidence. Given the population-based optimization method (PSO), a statistical evaluation of the swarm history in the optimization is performed. Usually, the likelihood approach is employed as it considers the system's nonlinearities. This approach can be employed using the Fisher−Snedecor criterion. All details of the deduction of this criterion for single objective process optimization are beyond the scope of this work. It is presented elsewhere.[19] Here, the intuition and some mathematical basis is presented.

The key idea is that the Lagrangian ($L$), which is the sum of the objective function with weighted vector of constraints ($C$), can be treated as a random variable, and the decision vector points are statistically compared using the Fisher−Snedecor. The Lagrangian and Fisher−Snedecor can be expressed, respectively, as

$$L(\boldsymbol{\theta}, \lambda) = F_{obj}(\boldsymbol{\theta}) + \boldsymbol{\lambda}^T \boldsymbol{C}(\boldsymbol{\theta}) \quad (2)$$



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

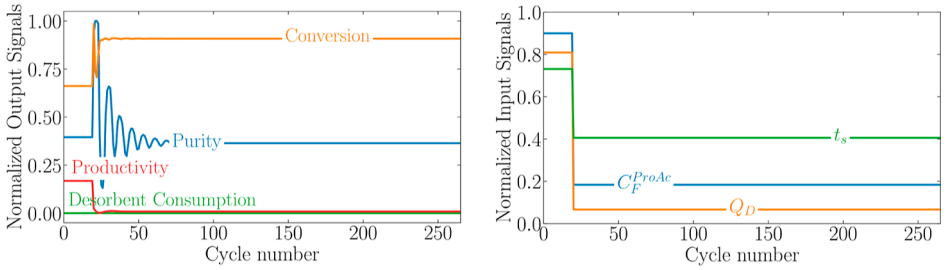



**Figure 4.** Input and output signals during experiment 13.

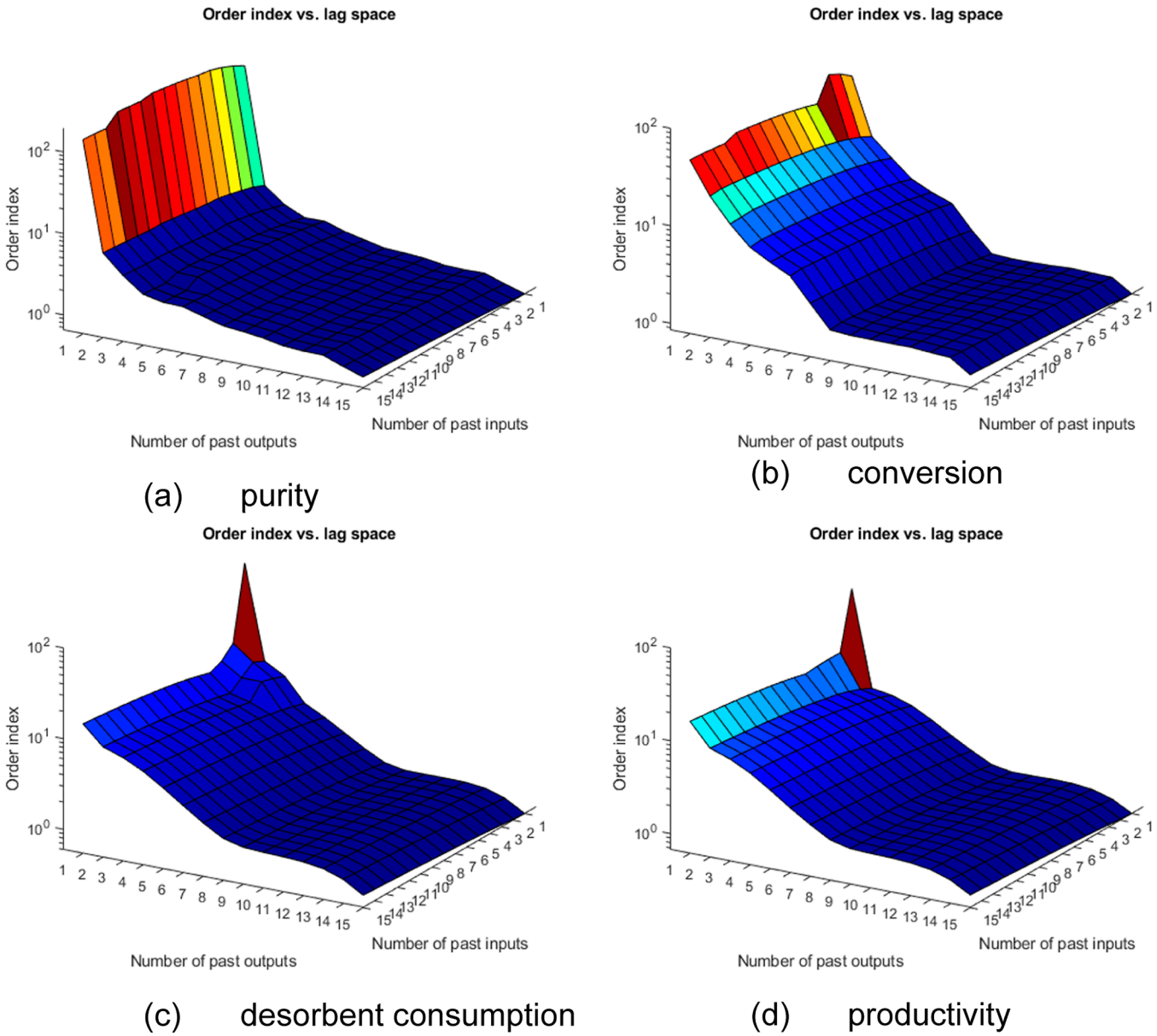



**Figure 5.** Lipschitz coefficients results for all performance indicators—(a) purity, (b) conversion, (c) desorbent consumption, and (d) productivity.

$$L(\boldsymbol{\theta}) \leq L(\boldsymbol{\theta}^*) + \frac{n_k n_\theta}{n_k - n_\theta - n_y + 1} F_\alpha(n_\theta, n_k - n_\theta - n_y + 1) \quad (3)$$

it compares the Lagrangian $L(\theta^*)$ at the minimum found in the optimization, $\theta^*$, with a given decision vector $\theta$ with objective function $L(\theta)$ using the Fisher−Snedecor distribution $F_\alpha$ with confidence level $\alpha$ and $n_k - n_\theta - n_y + 1$ degrees of freedom, where $n_y$ is the total number of performance indicators, $n_k$ is the number of iterations in the optimization, and $n_\theta$ is the number of operating variables (decision vector). Thus, the history of swarm position and objective function can be evaluated to compose an $\alpha$ confidence-level region when passing the inequality criteria described in inequality (3).

Equation 3 derivation is adapted from the confidence region evaluation for parameter estimation presented by Schwaab et



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

al. (2008)[38] and Benyahia et al. (2013).[41] In the proposal of Nogueira et al. (2019)[39] and here adapted, the methodology was adapted in order to evaluate the feasible operating region of the process operating variables after the optimization.

In Figure 3, $\epsilon$ represents the distance between the optimum analytical point and the optimum point found by the PSO. This concept allows us to determine the normalized squared error $(e_i^2)$ concerning the variance $(V_i)$, according to the following equation

$$e_i^2 = \frac{(L(\boldsymbol{\theta}_i^*, \lambda_i^*) - L(\boldsymbol{\theta}_i^g, \lambda_i^g))^2}{V_i}, \qquad \forall\, i = 1,2, \ldots, n_i \tag{4}$$

where $n_i$ is the number of decision variables of the PSO algorithm.

The error can be defined for all instances, $n_k$, obtained in the optimization process, as follows

$$e^2 = \sum_{i=1}^{n_i}\sum_{k=1}^{n_k} e_{ik}^2 = \sum_{i=1}^{n_i}\sum_{k=1}^{n_k} \frac{(L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) - L(\boldsymbol{\theta}_i^g, \lambda_{ik}^g))^2}{V_i} \tag{5}$$

the variance $V_i\ (\boldsymbol{\theta}_{ik}, \lambda_{ik})$ between the global optimum and the total number of points considered in optimization $L(\boldsymbol{\theta}_{ik}, \lambda_{ik})$ is defined as

$$\boldsymbol{V}_i(\boldsymbol{\theta}_{ik}, \lambda_{ik}) = \frac{1}{n_k}\left(\sum_{k=1}^{n_k} (L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) - L(\boldsymbol{\theta}_{ik}, \lambda_{ik}))^2\right) \tag{6}$$

combining 5 and 6

$$e^2 = \sum_{i=1}^{n_i}\sum_{k=1}^{n_k} \frac{(n_k)V_i(\boldsymbol{\theta}_{ik}, \lambda_{ik})}{V_i} \tag{7}$$

at this point, a chi-squared distribution for $e^2$, $\chi^2$, is assumed. This probability distribution has $n_k - (n_y - 1)$ degrees of freedom, where $n_k$ is the total number of iterations developed by the PSO and $n_y$ is the total number of objective functions considered in the optimization.

$$e^2 \to \chi^2(n_k - n_y + 1) \tag{8}$$

The error is calculated between the global analytical optimum and the remaining points belonging to the PSO population, represented in Figure 3, by $h$. The mathematical expression for $h$ is defined as

$$h_i^2 = \sum_{k=1}^{n_k-1} \frac{(L(\theta_{ik}^*, \lambda_{ik}^*) - L_i(\boldsymbol{\theta}_{i(k+\Delta k)}, \lambda_{i(k+\Delta k)}))^2}{V_{ik}} \tag{9}$$

extending the determination of error $h$ to all points in each evaluation of the objective function about the analytic global optimum, we have the following normalized expression

$$h^2 = \sum_{i=1}^{n_i} h_i^2 = \sum_{i=1}^{n_i}\sum_{k=1}^{n_k} \frac{(n_k)V_{ik}(\boldsymbol{\theta}_{i(k+\Delta k)}, \lambda_{i(k+\Delta k)})}{V_i} \tag{10}$$

where the variance $V_{ij}$ can be defined by

$$V_{ij}(\boldsymbol{\theta}_{i(k+\Delta k)}, \lambda_{i(k+\Delta k)}) = \frac{1}{n_k}\left(\sum_{k=1}^{n_k-1} (L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) - L(\boldsymbol{\theta}_{i(k+\Delta k)}, \lambda_{i(k+\Delta k)}))^2\right) \tag{11}$$

therefore, the assumed chi-square distribution premise for $h^2$ was also assumed for $e^2$, whose degrees of freedom are expressed by $n_k - (n_i + n_y - 1)$, as follows

$$h^2 \to \chi^2(n_k - n_i - n_y + 1) \tag{12}$$

thus, differences between the errors presented can be expressed by

$$h^2 - e^2 \to \chi^2(n_j) \tag{13}$$

these distributions are independent; it becomes a Fisher–Snedecor distribution as

$$\frac{\frac{e^2 - h^2}{n_i}}{\frac{h^2}{n_k - n_\theta + 1}} \to F_\alpha(n_i, n_k - n_\theta + 1) \tag{14}$$

where $\alpha$ is the confidence level of the Fisher–Snedecor test.

Finally, a Taylor series expansion can be derived around the optimal point to express the objective function

$$L(\boldsymbol{\theta}_{ik}, \lambda) = L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) + (\boldsymbol{\theta}_{ik} - \boldsymbol{\theta}_{ik}^*)\nabla \boldsymbol{L}_{\theta^*} + \frac{1}{2}(\boldsymbol{\theta}_{ik} - \boldsymbol{\theta}_{ik}^*)^T \boldsymbol{H}_{\theta^*}(\boldsymbol{\theta}_{ik} - \boldsymbol{\theta}_{ik}^*) \tag{15}$$

where $\nabla \boldsymbol{L}_{\theta^*}$ is the gradient vector and $\boldsymbol{H}_{\theta^*}$ is the Hessian matrix of the objective function, which is correlated with the covariance matrix optimal points as

$$\boldsymbol{H}_{\theta_{ik}^*} = 2V_{\theta_{ik}}^{-1} \tag{16}$$

replacing eq 14 in 13, one obtains

$$L(\boldsymbol{\theta}_{ik}, \lambda_{ik}) - L_i(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) = (\boldsymbol{\theta}_{ik} - \boldsymbol{\theta}_{ik}^*)^T \boldsymbol{V}_{\theta_{ik}}^{-1}(\boldsymbol{\theta}_{ik} - \boldsymbol{\theta}_{ik}^*), \equiv \chi^2(n_i) \tag{17}$$

simplifying eq 12, it is possible to rewrite it by equaling eq 15, obtaining

$$L(\boldsymbol{\theta}_{ik}, \lambda_{ik}) - L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) \to \sum_{i=1}^{n_i} \frac{(n_k)V_{ik}(\boldsymbol{\theta}_{ik}, \lambda_{ik})}{V_i(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*)} \frac{n_i}{n_k - n_\theta + 1} F_\alpha (n_i, n_k - n_\theta + 1) \tag{18}$$

assuming that $V_{ij}$ is a good approximation for $V$, the optimal region meets the following test

$$L(\boldsymbol{\theta}_{ik}, \lambda_{ik}) \le L(\boldsymbol{\theta}_{ik}^*, \lambda_{ik}^*) + \frac{n_k n_i}{n_k - n_\theta + 1} F_\alpha (n_j, n_k - n_\theta + 1) \tag{19}$$

## 3. RESULTS AND DISCUSSION

**3.1. Identification.** A set of 800 steps with ranges shown in Table 1 were created with LHS. Note that each step persisted for 250 time steps (prior estimated system settling time), leading to a total of 250.000 points. Three variables could be subject to step changes—concentration of propanoic acid in the feed stream ($C_F^{ProAc}$), desorbent flow rate ($Q_D$), and switch time ($t_s$), and four performance indicators outputs were collected—time-averaged conversion, productivity, desorbent consumption, and raffinate purity.



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

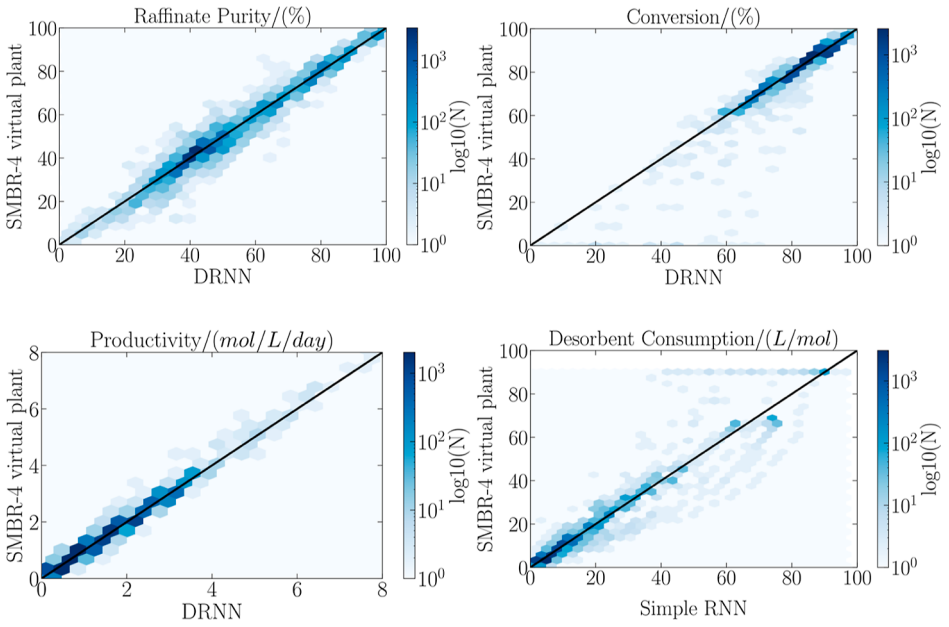



**Figure 6.** Hexabin parity plots for all performance indicators on the test set for the DRNN.

The remaining input variables and parameters necessary to simulate the SMBR-4 model were taken from, namely, feed stream flow rate ($Q_F$), recycling stream flow rate ($Q_{IV}$), and extract stream flow rates ($Q_X$), whose values are $1.400 \times 10^{-4}$, $1.2135 \times 10^{-4}$, and $9.000 \times 10^{-4}$ (in $dm^3/min$), respectively. Table 1 shows the ranges of the designed inputs as well as the outputs. The input space ranges were selected according to the physical limits of a lab-scale SMBR unit and made large enough to collect a considerable amount of information. It is worth remarking that the concentration of the limiting reactant cannot surpass its molar volume at working temperature, and it was the upper limit value for $C_F^{ProAc}$.

As described in 1.2, the original data set was partitioned into "experiments". Figure 3 shows a sample of a single experiment described for both input and output signals. Note that all variables are already rescaled. It is possible to note that the performance indicators have very distinct dynamic behavior, which justifies the MISO approach. 680 patterns (85% of first data set) similar to the one presented in Figure 3 were used for training and the remaining ones were used for hyperparameter tuning. A separate data set was generated and used for testing.

As described, the Lipschitz coefficient method was used to determine the order, that is, assess long-term dependencies of the system encoded in $n_a$ and $n_b$ parameters. These numbers were determined using a graphical approach. This analysis was run for each performance indicator as it deals with a MISO case. The orders are defined as the pair ($n_a$, $n_b$) at which the order index stops changing significantly as the order increases. From Figure 3, it can be seen that the purity, conversion, desorbent consumption, and productivity orders are (4,1), (8,1), (9,1), and (9,1), respectively.

In order to find a combination of hyperparameters that result in the minimum mean squared error of the validation set, HYPERBAND parameter maximum epochs were set to 200. The factor (proportion of discarded configurations) was set to 5. The HYPERBAND method was used with early-stopping regularization to prevent overfitting. Additionally, the proposed optimality assessment is a robust method for demonstrating generalization capacity.

Table 2 shows the search space of the hyperparameters for the RNN.

Table 3 shows the best found hyperparameters for each performance indicator of the RNN. It can be seen that for some performance indicators, the best architectures involves stacked LSTM and GRU cells. A similar pattern is observed for the activation function.

These architectures achieve satisfactory performance for all performance indicators: 0.0243 for productivity, 0.209 for conversion, and 0.710 for purity (considering mean absolute error for the whole validation set). Figure 5 shows hexabin parity plots for all performance indicators for the DRNN in the test set. On the *x*-axis are the DRNN predictions, and on the *y*-axis are the test data. The dynamic data and predictions are also presented in Figure 6 for the test set.

One important aspect to consider in black-box models is the accuracy/computational time trade-off. While the gPROMs simulator takes about 566 s to run the response for one single step perturbation, the RNN surrogate takes about 150 $\times 10^{-3}$ seconds (both running in an Intel i7-7500U—2.7 GHz CPU). As is demonstrated in the next sections, the simulation time difference becomes essential for optimization.

**3.2. Optimization for 95% Raffinate Purity Requirement.** The PSO algorithm was used to solve the optimization problem described in eq 1 with both the rigorous and DRNN models. The decision vector is $x = [\overline{t}_s, \overline{Q}_d, \overline{C}_F^{ProAc}]$, where the variables are the scaled switching time, desorbent flow rate, and concentration of propanoic acid in the feed stream, respectively, according to Table 1. They are the same ones used to train the DRNN. Thus, the intervals were large enough to cover several possible operating conditions. However, expanding these intervals beyond the input space where the DRNN was trained can lead to inaccurate calculations of the objective function. It is a drawback of all surrogate models; that is, they can only be made accurate in the identification region. However, this work generated data in a broad region using LHS based on the constraints of a real lab-scale plant. Therefore, it ensures that the identified models will represent the system under evaluation well.



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

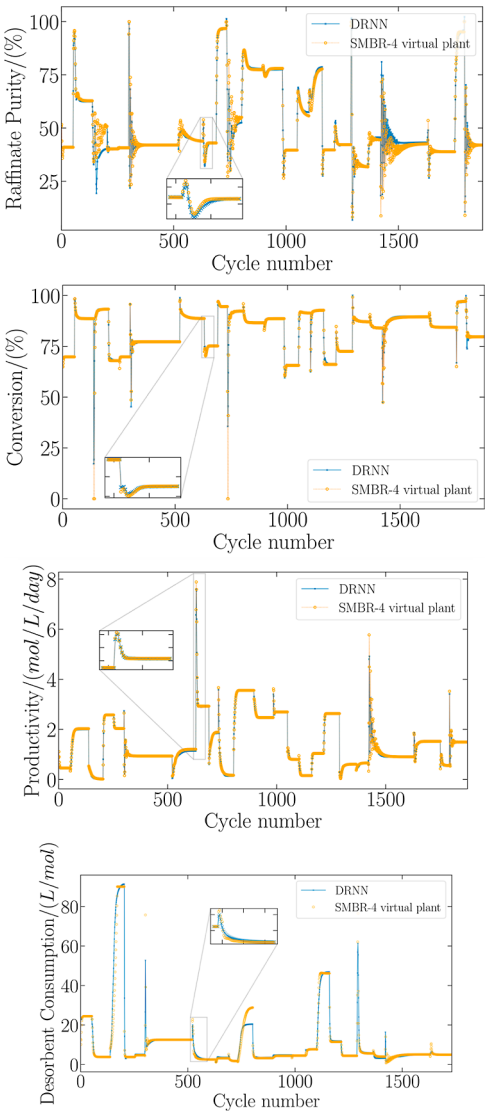



**Figure 7.** Comparison of DRNN predictions with a virtual SMBR-4 plant in the test set.

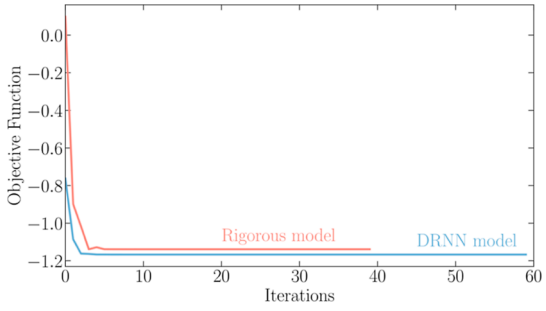



**Figure 8.** Evolution of swarm global best for 95% raffinate purity requirement.

$w_{conv} = -4 \times 10^{-3}$, $w_{ec} = 3 \times 10^{-1}$, $w_{prod} = -3 \times 10^{-1}$, and $w_{pur} = 50$ were chosen as the set of weights in the objective function, and $Pur_{min}$ was set to 95%. The weights reflect the objective from a process point of view—a negative value means that increasing the corresponding performance indicator magnitude will favor the objective function minimization and vice versa. Since the variables were rescaled, the upper and lower limits are 0 and 1, respectively.

When the SMBR-4 rigorous model was used in the optimization, 50 particles and 40 iterations were set. When the DRNN was used, 400 particles and 400 iterations were set. As mentioned in Section 2.1, the rigorous model is very costly to simulate. The number of particles and iterations is limited to the available time to perform the optimization. This specific case took 56 h with the rigorous model and 0.75 h with the DRNN.

Figure 4 shows the swarm minimum values' (global best) evolution over the iterations. Even though the optimization of the DRNN was carried out up to 400 iterations, only 60 are shown. It can be seen that the convergence of the optimization with the DRNN model is quicker. Still, both objective functions stop decreasing before the 10th iteration.

Table 4 shows the decision variables at the optimal point for both rigorous and DRNN models. It can be seen that the decision vectors at the optimal point are close to each other, and the highest discrepancy is found for the desorbent flow rate component ($Q_D$).

Table 5 shows the values of performance indicators that compose the objective function at the optimal point for each model and are evaluated using the same model; that is, DRNN optimal is evaluated with the DRNN model and rigorous optimal is evaluated with the rigorous model. It can be seen that the output values are close to each other. However, using single points to compare the optimization results of the two models can lead to misleading conclusions. One of the advantages of PSO is that it is a population-based algorithm which allows the building of the feasible operating region (FOR).

The DRNN optimal point could be run in the rigorous model, leading to an important perspective about the surrogate model optimal point robustness. However, the operating region characterization is a more complete and rigorous way of analyzing optimality. Furthermore, the proposed strategy is meant to demonstrate that the ANN result is equivalent to the rigorous model results. If this demonstration is done, it is proven that the optimizations are equivalent.

Figures 5 and 6 depict a 3D representation of swarm particles across iterations that meet the minimum purity requirement for both the rigorous model and DRNN, respectively. The *x*- and *y*-axis represent the scaled decision variables, and the *z*-axis is the objective function. The markers are colored according to the objective function value with a legend indicated in the color bar. In Figures 5b, 6b, 5c, and 6c, it can be seen that most of the swarm positions lie in a line where the concentration of the limiting reactant is close to the scaled upper limit 1. This value is a physical limit since the concentration of the component in the feed stream cannot surpass its molar density at a given temperature and pressure. Propanol feed concentration was calculated using its molar density for each value of propanoic acid concentration; that is, it is linearly correlated with $C_F^{ProAc}$ and removed from the inputs to avoid input space collinearity.

*3.2.1. Feasible Operating Region.* The points obtained from the optimization with a rigorous model and the surrogate DRNN model are used to build their corresponding confidence regions. A more reliable way to verify if the two models lead to the same optimal point than simply comparing the global best is by comparing the two confidence regions. Suppose a significant degree of superposition of the two regions is verified. In that case, the surrogate model can be


https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

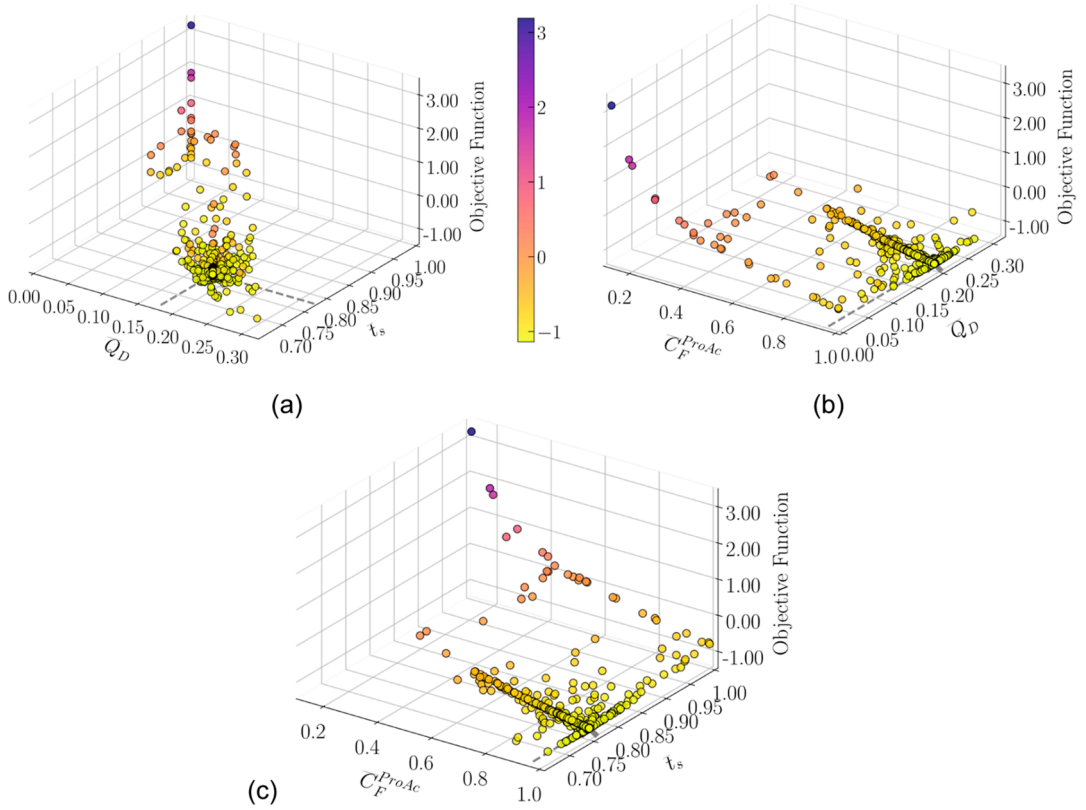



**Figure 9.** Swarm position and objective function for the rigorous model. (a) Desorbent flow rate vs switching time view, (b) desorbent flow rate vs limiting reactant feed concentration view, and (c) switching time vs limiting reactant feed concentration view.

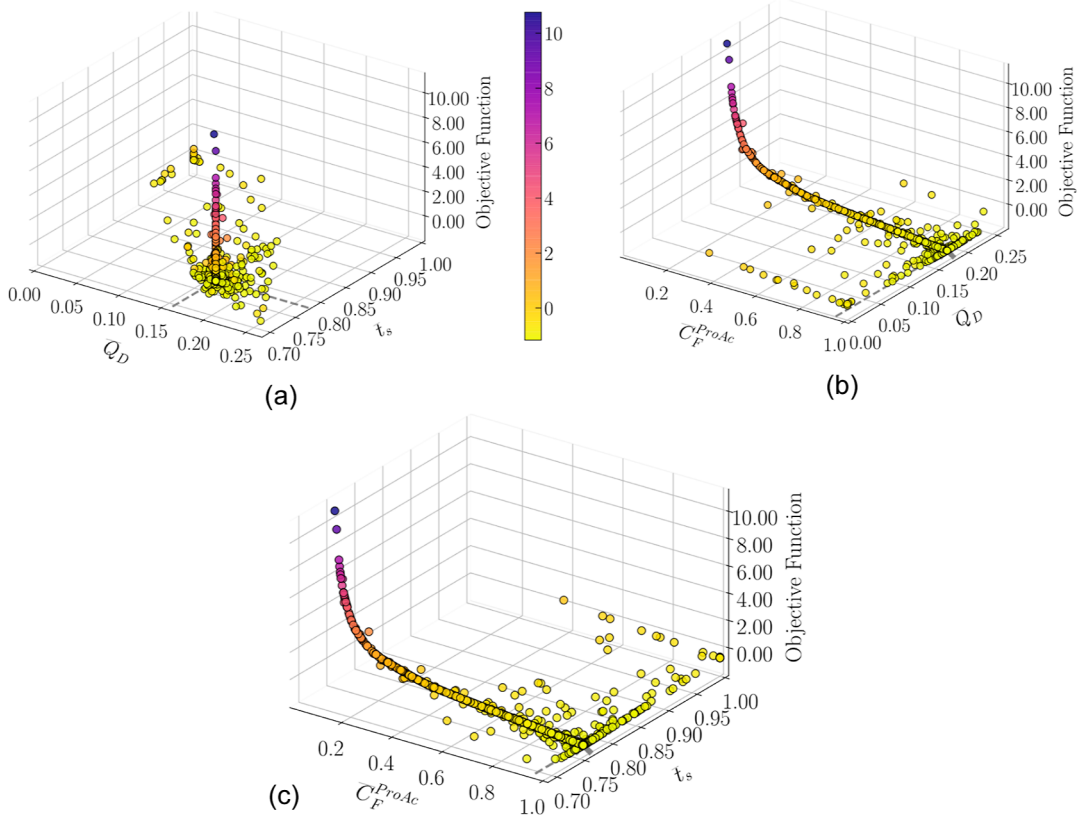



**Figure 10.** Swarm position and objective function for the DRNN model. (a) Desorbent flow rate vs switching time view, (b) desorbent flow rate vs limiting reactant feed concentration view, and (c) switching time vs limiting reactant feed concentration view.



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

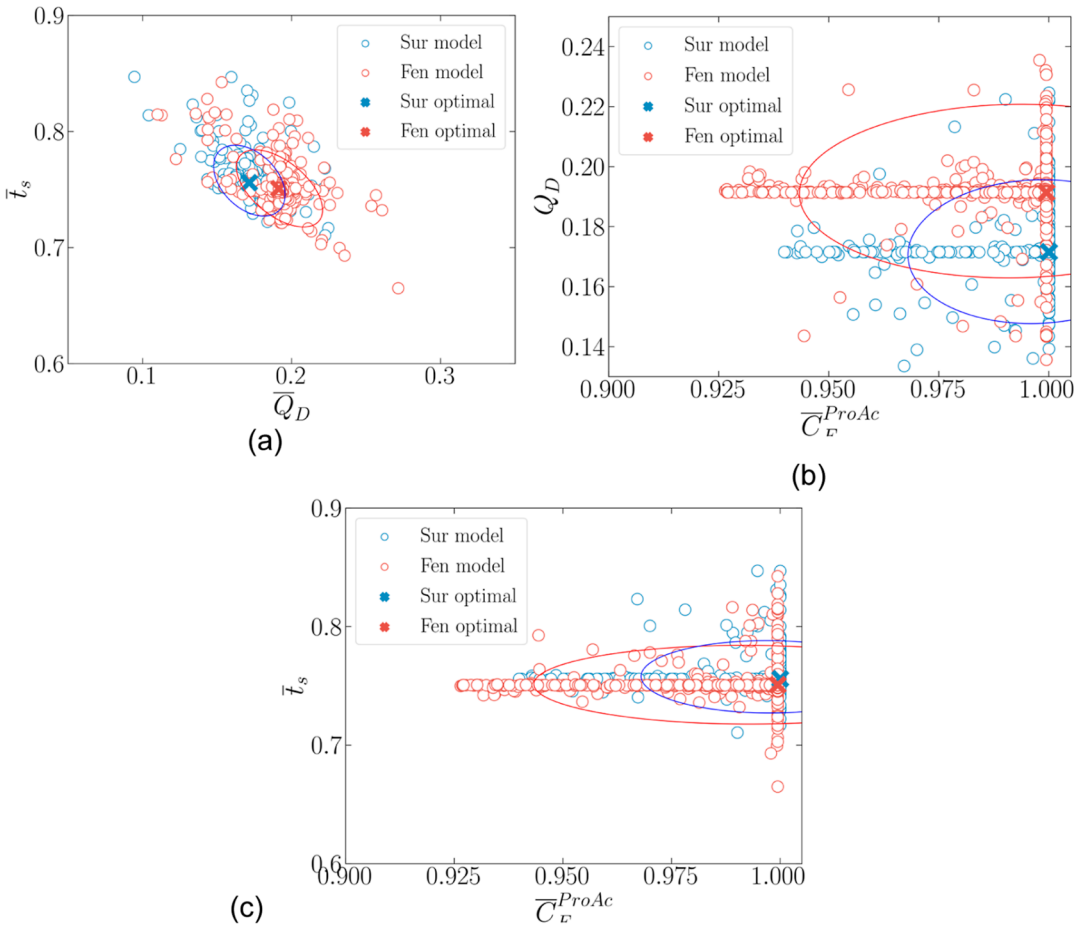



**Figure 11.** Scatter plot of points for 99.5% confidence level for rigorous (Fen) and surrogate (Sur) model (DRNN). (a) Desorbent flow rate vs switching time view, (b) limiting reactant feed concentration vs desorbent flow rate view, and (c) limiting reactant feed concentration vs switching time view.

reliably used to optimize searches within the domain it was identified.

Figure 7 shows a pair-wise scatter plot of the decision variables that passed a 99.5% test. It also includes the 99.7% probability ellipse (3 standard deviations) fitted using three bivariate Gaussian density distributions (one for each pair of inputs). In Figure 7a,b, it can be seen that most of the points lie in two lines of either constant $\bar{C}_F^{ProAc}$ or $\bar{Q}_D$ and $\bar{t}_s$, respectively, indicating that the Gaussian density may not be the ideal probability density function to fit. In Figure 7c, the Gaussian density seems to fit well in both clusters. Comparing the rigorous model with the DRNN, it can be seen that the points show a similar distribution pattern around their optimal points, and the ellipses have a significant degree of superposition. Moreover, it is possible to observe that the desorbent flow rate and the switching time have an inverse correlation, that is, less desorbent can be used, and similar performance can be achieved as soon as the switching time is increased (Figures 8−11).

## 4. CONCLUSIONS

The SMBR is challenging to optimize through first-principles models. In this sense, previous works in the literature concerning surrogates for its simpler version, the SMB, and other cyclic adsorptive processes indicate that using surrogate models may offer several benefits for SMBR optimization. However, there are no works published in the literature proposing surrogate models for the SMBR to our knowledge. Moreover, when surrogate models are used for optimization, optimality issues are barely considered; that is, only the optimal point found is tested in the rigorous model and used as the performance metric. It may lead to misleading conclusions about the model quality for optimization. In this setting, the present work shows that the proposed framework for SMBR units' optimization based on DNN models and PSO optimizer with optimality evaluation provides a robust tool for SMBR surrogate identification and addresses optimality issues through the process feasible operating region.

A single objective function optimization problem was formulated by weighting the SMBR performance metrics. With this setting, a 448× speed up in the optimization time was observed when using the surrogate model compared to the rigorous SMBR model, with a significant overlap of the FOR regions for all decision variables. It is worth mentioning that this speed up is calculated to compare run times for online applications where both models are already identified and the goal is obtaining the optimal solution. Rigorous real-time optimization for such a system would be computationally infeasible due to the associated computational effort. The speed up is calculated by the ratio between mechanistic model simulation (1 dynamic response) elapsed time and surrogate model simulation elapsed time for the same response and in the same computer. Even sacrificing physics, the surrogate



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

## Table A1. SMBR—4 Model Equations

Material balance in volume elements of column $k$ for component $i$ and velocity variation equation:

$$\frac{\partial C_{i,k}}{\partial t} + \frac{\partial}{\partial x}(u_k C_{i,k}) + \frac{1-\varepsilon_b}{\varepsilon_b} K^L_{i,k}(q^*_{i,k} - q_{i,k}) = D_{ax,k}\frac{\partial^2 C_{i,k}}{\partial x^2},$$

$$\frac{\partial q_{i,k}}{\partial t} = \frac{3}{R_p} K^L_{i,k}(q^*_{i,k} - q_{i,k}) + \vartheta_i \frac{\rho_b}{(1-\varepsilon_b)} r_{ProPro}$$

$$\frac{\partial u_k}{\partial x} = -\frac{1-\varepsilon_b}{\varepsilon_b}\frac{3}{R_p}\sum_{i=1}^{N_c} K^L_{i,k} V^M_i (q^*_{i,k} - q_{i,k}) + \frac{\rho_b}{1-\varepsilon_b} r_{ProPro}\sum_{i=1}^{N_c}\vartheta_i V^M_i,$$

$N_C$ is the total number of components, $x$ is the axial position, $t$ is the time, $\vartheta$ is the stoichiometric coefficient, $r_{ProPro}$ is the reaction rate, $\rho_b$ is the bulk density, $D_{ax}$ is the axial dispersion coefficient, $q^*$ is the adsorbed molar concentration in equilibrium with $C$ (bulk phase molar concentration), $q$ is the average adsorbed molar concentration, $u$ is the liquid interstitial velocity, $R_P$ is the mean particle radius, $\varepsilon_b$ is the bed porosity, $K^L$ is the external mass transfer coefficient, and $V^M$ is the molar volume.

Initial and boundary conditions for each component $i$ and column $k$:

$$t = 0, \quad \forall\, x \in\, ]0, L_k[ \rightarrow C_{i,k}(t,x) = C^{initial}_{i,k};\ q^{initial}_{i,k} = q^*_{i,k}(C^{initial}_{i,k})$$

$$x = L_k^{in}, \forall\, t \rightarrow u(0)_k C_{i,k} - D_{ax,k}\frac{\partial C(0,t)_{i,k}}{\partial x} = u(0)_k C^{in}_{i,k},$$

$$x = L_k^{out}, \forall\, t \rightarrow \frac{\partial C(L_k,t)_{i,k}}{\partial x} = 0$$

$L_k^{in}$ and $L_k^{out}$ are the inlet and outlet positions in section $k$, respectively

Global and component balances in nodes:

desorbent node:

$$u^{out}_{IV} + u_D = u^{in}_I,$$

$$C^{out}_{i,IV}u^{out}_{IV} + C_{i,d}u_d = C^{in}_{i,I}u^{in}_I$$

feed node:

$$u^{out}_{II} + u_F = u^{in}_{III},$$

$$C^{out}_{i,II}u^{out}_{II} + C_{i,f}u_f = C^{in}_{i,III}u^{in}_{III}$$

extract node:

$$u^{in}_{II} + u_X = u^{out}_I,$$

$$C^{in}_{i,II} = C^{out}_{i,I} = C_{i,X}$$

raffinate node:

$$u^{in}_{IV} + u_R = u^{out}_{III},$$

$$C^{in}_{i,II} = C^{out}_{i,I} = C_{i,X}$$

intermediate nodes:

$$C^{in}_{i,k+1} = C^{out}_{i,k};$$

$$u^{in}_k = u^{out}_{k+1}$$

Kinetic and equilibrium equations:

*isotherm model:*

$q^*_i = \frac{Q_i^{max} K_i C_i}{1 + \sum_{i=1}^{N_c} K_i c_i}$ $Q_i^{max}$ is the maximum adsorption capacity, $K_i$ is the adsorption constant

*Chemical reaction equation:*

$ProAc_{(l)} + POH_{(l)} \leftrightarrows H_2O_{(l)} + ProPro_{(l)}$; ProAc is propanoic acid, POH is 1-propanol, $H_2O$ is water, and ProPro is *n*-propyl-propionate

*Reaction rate model:*

$r_{ProPro} = 1.535 \times 10^{-4}\left(a_{ProAc}a_{POH} - \frac{a_{H_2O}a_{ProPro}}{37.09}\right)$, $a$ is the activity coefficient

*Mass transfer coefficient correlations:*

$$Sh_{p,i} = \frac{K^L_i d_p}{D_{i,mix}}$$

$$Re_p = \frac{\rho u d_p}{\eta_{mix}}$$

$$D_{i,mix}\eta_{mix}^{0.8} = \sum_{\substack{j=1 \\ i\neq j}}^{n} x_j D^0_{i,j}\eta_j^{0.8}$$

$$Sh_{p,i} = 1.09(Re_p Sc)^{0.33}$$

$$Sc = \frac{\eta_{mix}}{\rho D_{i,mix}}$$

$$D^0_{i,j} = \frac{8.2\times 10^{-8}}{\eta_j V_{M,i}^{1/3}}\left[1 + \left(\frac{3V_{M,j}}{V_{M,i}}\right)^{2/3}\right]$$

$Sh_p$, $Re_P$, and $Sc$ are the Sherwood, Reynolds, and Schmidt numbers, respectively, $d_p$ is the mean particle diameter, $D_{i,mix}$ is the mixture diffusion coefficient, $\rho$ is the fluid density, $\eta_{mix}$ is the mixture viscosity, and $D_{i,j}^0$ is the infinity dilution diffusion coefficient.

Performance indicators:

conversion:

$$Conv = 1 - \frac{\int_t^{t+Nt^*}(u_X C^X_{ProAc} + u_R C^R_{ProAc})dt}{u_f C^f_{ProAc} Nt^*}\ (\%)$$

purity:

$$Pu_R = \frac{\int_t^{t+Nt^*} C^R_{ProPro} dt}{\int_t^{t+Nt^*}(C^R_{ProPro} + C^R_{ProAc} + C^R_{H_2O})dt}\ (\%)$$

productivity:

$$Prod = \frac{\varepsilon_b u_R \int_t^{t+Nt^*} C^R_{ProPro} dt}{(1-\varepsilon_b)L_{unit} Nt^*}\left(\frac{mol}{L\ day}\right)$$

desorbent consumption:

$$Dc = \frac{(C^d_{POH}u_d - u_f(C^f_{POH} - C^f_{ProAc}X))V_{mol.POH}}{C^R_{ProPro}u_R}\left(\frac{L_{POH}}{mol_{ProPro}}\right)$$

model is feasible for real-time usage, where the fully mechanistic model, despite being run in gPROMs with a top desktop processor, fails to achieve it. Moreover, the rigorous model used is not easily parallelizable and would not benefit significantly from a high-performance computing environment as it is a stiff system of partial differential equations, which does not have a steady state but a cyclic one.



https://doi.org/10.1021/acsomega.2c06737
ACS Omega 2023, 8, 6463−6475

This work built the surrogate model from noise-free data, that is, from rigorous numerical simulations of the real SMBR model. In future works, one may consider introducing computer-generated noise to simulated data and evaluating its effects on the model and the process optimization. However, the results obtained prove that the surrogate model represents the system behavior with precision. This is clearly seen by the optimality analysis. Finally, an optimal FOR was obtained. The SMBR can be operated with high conversion while providing high purity ProPro using the obtained FOR.

## APPENDIX

The model equations and parameters are displayed in Tables A1, A2, and A3.

**Table A2. Langmuir Multicomponent Competitive Model Parameters**

| component | $K$/(L/mol) | $Q_{max}$/(mol/$L_{ads}$) |
|---|---|---|
| POH | 11.66 | 9.13 |
| ProAc | 9.04 | 10.06 |
| ProPro | 5.08 | 5.11 |
| $H_2O$ | 2.35 | 43.07 |

**Table A3. Parameters of Columns in TMBR**

| parameter | value |
|---|---|
| $\varepsilon_b$ | $4 \times 10^{-1}$ |
| $P_e$ | 166 |
| $R_p$/($\mu$m) | 122.75 |
| $\rho_b/\left(\frac{g_{ads}}{dm^3_{bed}}\right)$ | 390 |
| $L$/(dm) | 4.6 |
| $A_c$/($dm^2$) | $4.15 \times 10^{-2}$ |
| $T$/(K) | 313 |

## AUTHOR INFORMATION


### Corresponding Authors

**Márcio A. F. Martins** − *Industrial Engineering Program, Polytechnic School, Federal University of Bahia, 40210-630 Salvador/Bahia, Brazil;* orcid.org/0000-0003-2390-1525; Email: marciomartins@ufba.br

**Idelfonso B. R. Nogueira** − *Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway;* orcid.org/0000-0002-0963-6449; Email: idelfonso.b.d.r.nogueira@ntnu.no

### Authors

**Vinícius V. Santana** − *Laboratory of Separation and Reaction Engineering, Associate Laboratory LSRE-LCM, Department of Chemical Engineering, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal; Industrial Engineering Program, Polytechnic School, Federal University of Bahia, 40210-630 Salvador/Bahia, Brazil;* orcid.org/0000-0002-2940-1268

**José M. Loureiro** − *Laboratory of Separation and Reaction Engineering, Associate Laboratory LSRE-LCM, Department of Chemical Engineering, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal*

**Ana M. Ribeiro** − *Laboratory of Separation and Reaction Engineering, Associate Laboratory LSRE-LCM, Department of Chemical Engineering, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal;* orcid.org/0000-0003-4269-1420

**Luana P. Queiroz** − *Laboratory of Separation and Reaction Engineering, Associate Laboratory LSRE-LCM, Department of Chemical Engineering, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal;* orcid.org/0000-0003-3092-8136

**Carine M. Rebello** − *Industrial Engineering Program, Polytechnic School, Federal University of Bahia, 40210-630 Salvador/Bahia, Brazil*

**Alírio E. Rodrigues** − *Laboratory of Separation and Reaction Engineering, Associate Laboratory LSRE-LCM, Department of Chemical Engineering, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal;* orcid.org/0000-0002-0715-4761

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c06737


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was financially supported by Brazilian research agency FAPESB under grant BOL0731/2020, EDUFI Fellowships program, LA/P/0045/2020 (ALiCE), UIDB/50020/2020, and UIDP/50020/2020 (LSRE-LCM), funded by national funds through FCT/MCTES (PIDDAC), FCT—Fundação para a Ciência e Tecnologia under CEEC Institutional program.

## REFERENCES


(1) Rodrigues, A. E.; Pereira, C.; Minceva, M.; Pais, L. S.; Ribeiro, A. M.; Ribeiro, A.; Silva, M.; Graça, N.; Santos, J. C. Simulated Moving Bed Reactor. *Simulated Moving Bed Technology*; Elsevier, 2015; pp 183−223.

(2) Migliorini, C.; Fillinger, M.; Mazzotti, M.; Morbidelli, M. Analysis of Simulated Moving-Bed Reactors. *Chem. Eng. Sci.* **1999**, *54*, 2475−2480.

(3) Minceva, M.; Gomes, P. S.; Meshko, V.; Rodrigues, A. E. Simulated Moving Bed Reactor for Isomerization and Separation of P-Xylene. *Chem. Eng. J.* **2008**, *140*, 305.

(4) Tie, S.; Sreedhar, B.; Agrawal, G.; Oh, J.; Donaldson, M.; Frank, T.; Schultz, A.; Bommarius, A.; Kawajiri, Y. Model-Based Design and Experimental Validation of Simulated Moving Bed Reactor for Production of Glycol Ether Ester. *Chem. Eng. J.* **2016**, *301*, 188−199.

(5) Nikačević, N. M.; Huesman, A. E. M.; Van den Hof, P. M. J.; Stankiewicz, A. I. Opportunities and Challenges for Process Control in Process Intensification. *Chem. Eng. Process.* **2012**, *52*, 1−15.

(6) Tie, S.; Sreedhar, B.; Donaldson, M.; Frank, T.; Schultz, A. K.; Bommarius, A. S.; Kawajiri, Y. Experimental Evaluation of Simulated Moving Bed Reactor for Transesterification Reaction Synthesis of Glycol Ether Ester. *Adsorption* **2019**, *25*, 795.

(7) Ray, N. M.; Ray, A. K. Multi-Objective Optimisation of Biodiesel Synthesis in Simulated Moving Bed Reactor. *Separations* **2021**, *8*, 127.

(8) Subramani, H. J.; Zhang, Z.; Hidajat, K.; Ray, A. K. Multiobjective Optimization of Simulated Moving Bed Reactor and Its Modification - Varicol Process. *Can. J. Chem. Eng.* **2004**, *82*, 590−598.

(9) Nogueira, I. B. R.; Viena, V.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M. Dynamics of a True Moving Bed Reactor: Synthesis of n-Propyl Propionate and an Alternative Optimization Method. *Chem. Eng. Process.* **2020**, *148*, 107821 No. January.




https://doi.org/10.1021/acsomega.2c06737
*ACS Omega* 2023, 8, 6463−6475

(10) Nogueira, I. B. R.; Ribeiro, A. M.; Rodrigues, A. E.; Loureiro, J. M. Dynamic response to process disturbances-A comparison between TMB/SMB models in transient regime. *Comput. Chem. Eng.* **2017**, *99*, 230−244.
(11) Santana, V. V.; Martins, M. A. F.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M.; Nogueira, I. B. R. Transient Analysis of True/Simulated Moving Bed Reactors: A Case Study on the Synthesis of n-Propyl Propionate. *Comput. Chem. Eng.* **2020**, *137*, 106820.
(12) Li, L.; Dai, C.; Deng, J.-B.; Yuan, D.-C. Parameter Optimization of Neural Network Model Based on Simulated Moving Bed. *2019 Chinese Automation Congress (CAC)*; IEEE, 2019; pp 2506−2509.
(13) Wang, C.; Klatt, K.; Dünnebier, D.; Engell, S.; Hanisch, F. Neural Network-Based Identification of SMB Chromatographic Processes. *Control Eng. Pract.* **2003**, *11*, 949−959.
(14) Nogueira, I. B. R.; Ribeiro, A. M.; Requião, R.; Pontes, K. V.; Koivisto, H.; Rodrigues, A. E.; Loureiro, J. M. A Quasi-Virtual Online Analyser Based on an Artificial Neural Networks and Offline Measurements to Predict Purities of Raffinate/Extract in Simulated Moving Bed Processes. *Appl. Soft Comput. J.* **2018**, *67*, 29−47.
(15) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436.
(16) Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Network.* **1989**, *2*, 359.
(17) Park, C. Determination of the Joint Confidence Region of the Optimal Operating Conditions in Robust Design by the Bootstrap Technique. *Int. J. Prod. Res.* **2013**, *51*, 4695−4703.
(18) Nogueira, I. B. R.; Martins, M. A. F.; Requião, R.; Oliveira, A. R.; Viena, V.; Koivisto, H.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M. Optimization of a True Moving Bed Unit and Determination of Its Feasible Operating Region Using a Novel Sliding Particle Swarm Optimization. *Comput. Ind. Eng.* **2019**, *135*, 368−381.
(19) Rebello, C. M.; Martins, M. A. F.; Loureiro, J. M.; Rodrigues, A. E.; Ribeiro, A. M.; Nogueira, I. B. R. From an Optimal Point to an Optimal Region : A Novel Methodology for Optimization of Multimodal Constrained Problems and a Novel Constrained Sliding Particle Swarm Optimization Strategy. *Mathematics* **2021**, *9*, 1808.
(20) Process Systems Enterprise. Next-Generation Modelling Tools across the Process Lifecycle https://www.psenterprise.com/products/gproms (accessed 15-12-2022).
(21) McKay, M. D.; Beckman, R. J.; Conover, W. J. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **1979**, *21*, 239−245.
(22) Niederreiter, H. *Random Number Generation and Quasi-Monte Carlo Methods*; Society for Industrial and Applied Mathematics, 1992.
(23) Almeida, E. C.; Rebello, C. de M.; Fontana, M.; Schnitman, L.; Nogueirados, I. B. R. A Robust Learning Methodology for Uncertainty-Aware Scientific Machine Learning Models. **2022**, arXiv 2022.10.48550/ARXIV.2209.01900.
(24) Rebello, C. M.; Marrocos, P. H.; Costa, E. A.; Santana, V. V.; Rodrigues, E.; Ribeiro, A. M.; Nogueira, I. B. R. Machine Learning-Based Dynamic Modeling for Process Engineering Applications : A Guideline for Simulation and Prediction from Perceptron to Deep Learning. *Processes* **2022**, *10*, 250.
(25) Santana, V. V.; Martins, M. A. F.; Loureiro, J. M.; Ribeiro, A. M.; Rodrigues, A. E.; Nogueira, I. B. R. Optimal Fragrances Formulation Using a Deep Learning Neural Network Architecture: A Novel Systematic Approach. *Comput. Chem. Eng.* **2021**, *150*, 107344.
(26) Miriyala, S. S.; Mittal, P.; Majumdar, S.; Mitra, K. Comparative Study of Surrogate Approaches While Optimizing Computationally Expensive Reaction Networks. *Chem. Eng. Sci.* **2016**, *140*, 44−61.
(27) Pantula, P. D.; Miriyala, S. S.; Mitra, K. KERNEL: Enabler to Build Smart Surrogates for Online Optimization and Knowledge Discovery. *Mater. Manuf. Process.* **2017**, *32*, 1162−1171.
(28) Miriyala, S. S.; Pujari, K. N.; Naik, S.; Mitra, K. Evolutionary Neural Architecture Search for Surrogate Models to Enable Optimization of Industrial Continuous Crystallization Process. *Powder Technol.* **2022**, *405*, 117527.
(29) Himmelblau, D. M. Accounts of Experiences in the APplication of Artificial Neural Networks in Chemical Engineering. *Ind. Eng. Chem. Res.* **2008**, *47*, 5782−5796.
(30) Marrocos, P. H.; Iwakiri, I. G. I.; Martins, M. A. F.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M.; Nogueira, I. B. R. A Long Short-Term Memory Based Quasi-Virtual Analyzer for Dynamic Real-Time Soft Sensing of a Simulated Moving Bed Unit. *Appl. Soft Comput.* **2022**, *116*, 108318.
(31) He, X.; Asada, H. A New Method for Identifying Orders of Input-Output Models for Nonlinear Dynamic Systems. *1993 American Control Conference*; IEEE, 1993; pp 2520−2523.
(32) Li, L.; Jamieson, K.; DeSalvo, G.; Rostamizadeh, A.; Talwalkar, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *J. Mach. Learn. Res.* **2018**, *18*, 1−52.
(33) Inapakurthi, R. K.; Miriyala, S. S.; Mitra, K. Deep Learning Based Dynamic Behavior Modelling and Prediction of Particulate Matter in Air. *Chem. Eng. J.* **2021**, *426*, 131221.
(34) Inapakurthi, R. K.; Miriyala, S. S.; Mitra, K. Recurrent Neural Networks Based Modelling of Industrial Grinding Operation. *Chem. Eng. Sci.* **2020**, *219*, 115585.
(35) Mogilicharla, A.; Mittal, P.; Majumdar, S.; Mitra, K. Kriging Surrogate Based Multi-Objective Optimization of Bulk Vinyl Acetate Polymerization with Branching. *Mater. Manuf. Process.* **2015**, *30*, 394−402.
(36) Miriyala, S. S.; Subramanian, V. R.; Mitra, K. TRANSFORM-ANN for Online Optimization of Complex Industrial Processes: Casting Process as Case Study. *Eur. J. Oper. Res.* **2018**, *264*, 294−309.
(37) Kennedy, J.; Eberhart, R. 47-Particle Swarm Optimization Proceedings, IEEE International Conference. *Proceedings of ICNN'95—International Conference on Neural Networks*; IEEE, 1995; Vol. *11*(1), pp 111−117.
(38) Schwaab, M.; Biscaia, E. C., Jr.; Monteiro, J. L.; Pinto, J. C. Nonlinear Parameter Estimation through Particle Swarm Optimization. *Chem. Eng. Sci.* **2008**, *63*, 1542−1552.
(39) Nogueira, I. B. R.; Martins, M. A. F.; Requião, R.; Oliveira, A. R.; Viena, V.; Koivisto, H.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M. Optimization of a True Moving Bed Unit and Determination of Its Feasible Operating Region Using a Novel Sliding Particle Swarm Optimization. *Comput. Ind. Eng.* **2019**, *135*, 368.
(40) Rebello, C. M.; Martins, M. A. F.; Rodrigues, A. E.; Loureiro, J. M.; Ribeiro, A. M.; Nogueira, I. B. R. A Novel Standpoint of Pressure Swing Adsorption Processes Multi-Objective Optimization: An Approach Based on Feasible Operation Region Mapping. *Chem. Eng. Res. Des.* **2022**, *178*, 590−601.
(41) Benyahia, B.; Latifi, M. A.; Fonteix, C.; Pla, F. Emulsion Copolymerization of Styrene and Butyl Acrylate in the Presence of a Chain Transfer Agent. Part 2: Parameters Estimability and Confidence Regions. *Chem. Eng. Sci.* **2013**, *90*, 110−118.


https://doi.org/10.1021/acsomega.2c06737
*ACS Omega* 2023, 8, 6463−6475